REMARKS

The above amendment is made in response to the Office Action mailed September 25, 2003. Claims 26-34 have been added. Claims 1-34 are pending in the present application. Claims 1-25 stand rejected. The Examiner's reconsideration is respectfully requested in view of the above amendment and the following remarks.

The Office Action advises Applicants to delete the underlined title in the abstract. The abstract has been amended. The Office Action also advises Applicants to "delete...the extraneous matter found at the bottom of the page. See MPEP 608.01(f)." It is unclear to applicants what the cited "extraneous matter" is; Applicants are also confused by the citation to MPEP § 608.01(f), which refers to the "Brief Description of the Drawings" section. Applicants respectfully request a clarification.

Applicants respectfully acknowledge that claims 7, 8, 10-15 and 22 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form. Claims 26-34 are the above-referenced claims rewritten in independent form. Accordingly, it is respectfully submitted that claims 26-34 are in condition for allowance.

Claims 1-6, 9, 16-21 and 23-25 were rejected under 35 U.S.C. § 103(a) as being unpatentable over Agrawal *et al.* (U.S. Patent No. 5,799,311) (hereinafter "Agrawal") in view of Ramaswamy *et al.* ("Efficient Algorithms for Mining Outliers from Large Data Sets – 05/2000") (hereinafter "Ramaswamy"). The rejection is respectfully traversed.

The Office Action argues that col. 3, lines 25-29 of Agrawal show "the input data set being biased in favor of the records of the non-target class," as claimed in claim 1. Agrawal does not show such an input data set.

The Office Action argues that col. 3, lines 40-43 of Agrawal disclose "constructing a decision tree from the input data set...based upon *multivariate subspace splitting criteria*." This is incorrect. Agrawal discloses constructing a decision tree using a "split test." The "split test" disclosed in Agrawal creates *univariate* decision trees, and not *multivariate* decision trees. This is shown explicitly Figures 1 and 2 of Agrawal. Figure 2 of Agrawal shows a "complex" *univariate* decision tree. For example, node 2 of Agrawal is split in a *single* attribute: age; node 3 is split in a *single* attribute: car type. This is specifically the type of *traditional* decision tree discussed on page 7, line 19 of the present Specification. Agrawal clearly does not show "constructing a decision tree from the input data set...based upon *multivariate subspace splitting criteria*," as claimed in claim 1.

The Office Action argues that col. 2, lines 44-46 of Agrawal disclose "classifying and *scoring* the records, based upon the decision tree and the *nearest* neighbor set of nodes." This argument poses at least two problems. First, the citation provided by the Office Action clearly does not show "*scoring* the records." Second, the Office Action admits that Agrawal does not disclose "computing distance functions for each of the leaf nodes," as claimed in claim 1. However, the "*nearest* neighbor set of nodes" cannot be determined without distance functions. As a result, Agrawal clearly does not show "classifying and *scoring* the records, based upon the decision tree and the *nearest* neighbor set of nodes," as claimed in claim 1.

As stated above, the Office Action admits that Agrawal does not disclose "computing distance functions for each of the leaf nodes." The Office Action further admits that Agrawal does not disclose "identifying, *with respect to the distance functions*, a nearest neighbor set of nodes for each of the leaf nodes based upon a respective closeness

of the nearest neighbor set of nodes to a target record of the target class." However, the Office Action cites col. 6, lines 52-54 as showing that "Agrawal discloses the use of 'a *nearest* neighbor set of nodes for each of the leaf nodes based upon a respective closeness of the *nearest* neighbor set of nodes to a target record of the target class.'" Col. 6, lines 52-54 of Agrawal state: "The records at each new leaf node are checked at block 23 to see if they are of the same class." This citation of Agrawal clearly does not show using "a *nearest* neighbor set of nodes for each of the leaf nodes based upon a respective closeness of the *nearest* neighbor set of nodes to a target record of the target class." Applicants reiterate that the *nearest* set of nodes cannot be determined without distance functions. This is claimed explicitly in claim 1: "identifying, *with respect to the distance functions*, a nearest neighbor set of nodes...." Further, Applicants assume that the Office Action's argument is based on it's broad latitude to interpret claims. However, Applicants respectfully submit that the Office Action's interpretation of the claims must be consistent with the interpretation of one skilled in the art, as well as consistent with the present Specification. The Office Action's interpretation that a determination that a new leaf node is of the same class, as disclosed in Agrawal, somehow shows using "a *nearest* neighbor set of nodes for each of the leaf nodes based upon a respective closeness of the *nearest* neighbor set of nodes to a target record of the target class, " as disclosed in claim 1, does not meet those criteria.

The Office Action argues that page 429, col. 2, lines 36-46 of Ramaswamy discloses "an analogous system [to Agrawal] that teaches the claimed features 'computing distance functions for each of the leaf nodes' as a means for using the square of the Euclidean distance as the distance metric." This is incorrect.

The first issue is whether Ramaswamy is analogous art to Agrawal. It clearly is not. Agrawal discloses a method and system for generating a decision tree classifier from at training set of records. On the other hand, Ramaswamy does not relate to generating a decision tree or even generally to decision tree classifiers. Instead, Ramaswamy relates specifically to efficient algorithms for *mining outliers*. That is, Ramaswamy is concerned only with finding "small patterns" (*i.e.*, outliers) from a set of data that is considerably dissimilar or inconsistent with the remainder of the data. Agrawal deals with an assumed training set. Therefore, the consideration of outliers is irrelevant to Agrawal. One of ordinary skill in the art would not combine the disparate teachings of Agrawal and Ramaswamy.

The second issue is whether the combination of Agrawal and Ramaswamy teach or suggest "computing distance functions for each of the leaf nodes," as claimed in claim 1. The combination clearly does not disclose such limitations. Ramaswamy discloses a distance function related to *outliers*. It was previously argued that Agrawal does not disclose the "nearest neighbor set of nodes." However, assuming, *arguendo*, that Agrawal discloses such limitations, the Office Action's arguments that the distance functions of Ramaswamy can be combined with Agrawal are without merit. The Office Action's proposed motivation that "providing a quicker computation time of identifying small patterns for a given data analysis" is irrelevant to Agrawal. Identifying small patterns (*i.e.*, outliers) is related solely to Ramaswamy. The Office Action's proposed motivation to "[allow] Agrawal's system the enhanced capability of improving the accuracy and the efficiency of the method for building [a] space splitting decision tree" is also irrelevant. More specifically, the motivation does not explain why the distance function used for

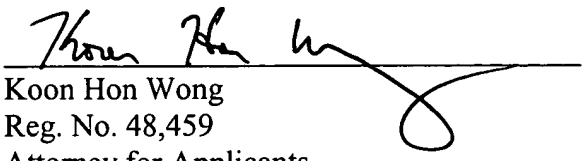
outliers, as disclosed in Ramaswamy, would somehow be applied to computing distance functions for each of the leaf nodes described in Agrawal. Agrawal does not disclose any use or need for such distance functions. A generic argument of improving accuracy and efficiency does not explain the combination of such disparate references. The Office Action's proposed motivation to "provide performance of the partition based algorithm relatively unchanged" is likewise irrelevant. More specifically, the comparison between index-based algorithm and the partition-based algorithm, each specifically for mining outliers, is entirely irrelevant to Agrawal or the problems it addresses. The Office Action's proposed modifications are, in effect, red herrings that do not produce a valid suggestion or motivation to combine the above-referenced teachings of Agrawal and Ramaswamy.

Applicants respectfully note that the Office Actions argument that the language in Ramaswamy is "*similar* to the description" provided by the Applicants is especially troublesome. A *similarity* in descriptions of a reference and the instant Specification is not *prima facie* evidence of obviousness. Such argument is irrelevant and should not be part of any basis of § 103(a) rejection.

Accordingly, claim 1 is believed to be patentably distinguishable over Agrawal in view of Ramaswamy. Independent claim 24 and dependent claims 2-23 are believed to be allowable for at least the reasons given for claim 1. Dependent claim 25 is believed to be allowable for at least the reasons given for claim 24. Withdrawal of the rejection of claims 1-25 is respectfully requested.

In view of the foregoing remarks, it is respectfully submitted that all the claims now pending in the application are in condition for allowance. Early and favorable reconsideration is respectfully requested.

Respectfully submitted,

By: 

Koon Hon Wong
Reg. No. 48,459
Attorney for Applicants

F. CHAU & ASSOCIATES, LLC
1900 Hempstead Turnpike, Suite 501
East Meadow, New York 11554
Telephone: (516) 357-0091
Facsimile: (516) 357-0092